

Figure 1

Anopheles gambiae arrestin 1 cDNA sequence (SEQ ID NO: 1)

5 ACAGGAACGACGGTTGTGATCCCTCACTGGTGGTACACGAATCATAAGCATTATTCATAACCT
AAAAAAACAAAATCTACAAAAAAAAGCTTCATTCCCATCGAAAAAAACTTCTTGTGAAATCAACCG
AGCTAACAAACAACATCCTGTGAAAATCTAGCAGTGAAGGTGTGATATCGTATACCTGTACCTG
TAAACCCTGTCGCGGTGTGCCTTGTGTATCAATTGTGGAAAACAGAAAATACATCAAAA
10 TGGTTACAATTCAAAGTCTCAAGAAGTGCCCCCTAATGGAAAGGTTACGCTGTACATGGG
CAAGCGTGACTTTGTAGACCACGTTCCGGCGTGAACCGATCGATGGTATCGTCGTCCCTCGAT
GATGAGTACATTCTGTGACAACCGTAAGGTATTGGTCAGATTGTCTGCAGTTCCGCTACGCC
GCGAAGAGGACGAGGTGATGGGACTAAACTTCCAGAAGGAGTTATGCCCTCGCTCCGAACAGAT
CTACCCCGCTCCGGAAAAGTCGGACAAGGAGCAGACCAAGCTCCAGGAGCGACTGCTGAAGAA
15 GCTGGGTTCGAACGCCATCCGTTCACGTTCAACATCTGCCGAATGCTCCGTCTCGGTACCG
CTGCAGCAGGGCGAACGATGATAATGGAGACCCGTGCGGTGTCTGACTACGTGAAGATCTTG
CCGGTGAGTCGGAAACCGATCGTACGCACCGTCGACGCCGTTACGCTCGGCATACGCAAGAT
CCAGTCGCACCGACAAGCAGGGCCAGCAGCCGTGACGCCGCTGGTCGCAAGGACTTATGCTA
AGCCCAGGAGAGCTGGAGCTCGAGGTACACTAGACAAGCAGCTGTACCTGCACGGGAGCGA
20 ATAGGCCTAACATCTGCATCCGCAACAACTCGAACAAAATGGTCAAGAAGATTAAAGGCCATGG
TCCAGCAGGGTGTGGATGTGGTGTCTCAGAATGGTAGTACCGCAACACAGTGGCATCGCT
GGAGACTAGCGAGGGTTGCCAATTCAAGCCCGCTCCAGTCTGCAGAAGGTAATGTACCTCACG
CCGCTGCTGCCTCGAACAAAGCAGCGACGTGGCATGCCCTGGACGGTCAGATCAAGCGTCAGG
ATCAGTGTGTTGGCCTCGACAAACCTCTGGCTCAACCGGATCAGCGAGATGCTTCCGGTTAT
25 CATATCGTATGCCGTAAGGTTAAGCTTTCTCGGCGACTCGGCCGAGCTGTCCGGAA
CTTCCATTGCTGATGCACCCAAAGGCCGACCAAGGCTAAGGTCAATGCCGACAGCC
AGGCCGACGTAGAAACTTCCGACAGGATACAATCGACCAGCAGGCATCAGTTGACTTGAATA
30 GACGACGCAACGGTTGGAAATGCTACCTACTACCCAGGCATGGCTAACACGACGAACGAAC
TACTACTACTAACGATAAAAACAGGAAAAAAATGGAAAACCTAAAAATGGATCATACAACCG
AACGCAAACGACCTACGACGATCGATCTCACTTCCCCTTTCATCCTAACGCAATAGAACGA
TGGTAGAAAAGGAAGATAAGATGGAGAGAAAGTCACGTGTATCAATGACGACGACTACCAAAA
35 CTGAAGACGTAACACATGTTCCCAGCGAGCGGTAACTGTTCTGACACCTCCGCTCGA
CAATGTTACCTTTAAAAACATACAAATTAGAAGTCGTCTCACTACCTCAACCAATCCAGGCCAC
TTGGTATATACTTTCATAGAATCCTCTGAGCGCAAGGACCTATTGAAATTCAAGTGTATT
GTAACTGCGACCAAATGCCCTAGCTGAATGTTGTGAACGAGTTATGTACATCAAAAGATTGAATA
AAACAAAAAAAAAAAAAA

40

Figure 2

Anopheles gambiae arrestin 1 amino acid sequence (SEQ ID NO: 2)

5

MVYNFKVFKKCAPNGKVTLYMGK RDFV D H VSGVEPIDGIVVLDDEYIRDNRKVFGQIVCSFRYGR
EEDEVMGLNFQKELCLASEQIYPRPEKSDKEQTKLQERLLKKLGSNAIPFTFNISPNA PSSVTLQQG
EDDNGDPCGVSY YVKIFAGESETDRTHRRSTVTLGIRKIQFAPTKQGQQPCTLVRKDFMLSPGELE
10 LEVTLDKQLYLHGERIGVNICIRNNSNK MVKKIKAMVQQGV DVVLFQNGSYRNTVASLETSEG CPI
QPGSSLQKV MYLTP LLSSNKQR RGIALDGQIKRQDQCLASTTLLAQP DQRDAFGVIISYAVKVKLFL
GALGGELSAELPFVLMHPKGTKAKVIHADSQADVETFRQDTIDQQASVD FE

15

Figure 3a

Anopheles gambiae odorant receptor 1 genomic sequence (SEQ ID NO: 9)

5

Features:

- 1) Presumed Untranslated 5' and 3' regions are underlined.
- 2) Potential TATA box transcription initiation signal is double underlined.
- 3) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
- 10 4) Introns are tentatively assigned and are shown in lower case.

Exons are **highlighted**.

AGCTTGTTCATTTATGAAATCTAGCCCATTGTATAGTCTAACGACGAAGAACATACGAAAGTACCTCGT
CCGAACACTATCAACATTAACTACCAAGCTAGAAGAAGATATTATAGTCAGCCTAACATCATAGGAAACTTT
15 AGAAAACCATTAACTTACATGATGATAAGTCCCACCTTACCCAGCACAGGTTGAGAAGGACGAAAGTATCT
TTACGATAATATTACTCTAAGGTAGTTTGAAATAAAATAAAATTACGTCAAGTGGCATCGGACATCATT
GAAAATCTACTAACGACATAACACACACCAAGACGACGACGTAGTTCATCTAGAAAAACGGTCAGCTCCATC
GAACACGTCAAGGACATAACTGCACATCGTATGGTCAGTCCACTAGTGCCAACACTGGTCCAGGGCACTACCTT
CCGAAGCAGTAGAACCTAATGTATTGAAATTATTAGGACATACTGCACATGCATATGGCTAGTTCCGCTGGTAC
20 AACGATGGCACAGGACACTATCTGGCCTTGTAAAATCACTGTAAAATCTATACAAAAACGGCTTACCCATACT
TTATCACAAAAACGGCAGGTGAGGGCTGGATTGTTCAAAGCATAGAAATATAATTCAAAGTCCATAATCTCC
TTAAAAGATAGACAaCAGTAGAGAACACATTAGTCTTCTGTTGAGTTAGTGCCTCTCAAGTAAGCGTT
AATGCTCAATTGTTGTAGATTGTTGGATGACTCTCGCTACGTCTATAGTGTCAATACTCCAATTAGATTCT
25 AATTAGTTCCAATTGTCACGGAAAACCCaCAAAGAAAAAAACTTGTATCTAGGGTGAATTTCGAGAACAA
ATTGGACACTTCATATGAAAAAGGACAGCTTTTCAAAATGTTAAATAAACACCGTTCGATCCTTgttgatttca
attctccaaattctgcagaataattctgcaaaattttacaactgctcaaccaccaataattccaattaatcatctg
aacatttaaaactgataattaagatgatgatattgttcgtcatcacctaagaatcgatttagttggataaaaagaa
caaattgaaataacaataaaagtccctgaattttattcgaataacggctgactcatttatttcaaaaacctttgaga
aattcctcgtaaaaattggctccatagttctgctaacggccacttcaaaagcaagaactaacaaaatcataat
30 tatggtgcagaactatcattaccgtatcgccattaaaaactttctcaatttgcggctgttaccggctaa
tacagagcagactacggaaagtgtatcaacgtcgctattgtataacgaggaaacgcctccgaaagggtgttgaagg
acctttcaattgaaacaagtactgtttccagttttaattggatagtataaaatgagccgttcaacgatcg
catcattgagttcatctcgaggagaaatagatcagtgccactgtttaaccgaaagtaatgaaagctgacaaact
gaaccacgggtggatcgatcgacgggattcggtcgactgttgcattttgaaatatttgc
35 GCCTATGGCCACCGGAAAGATA CGGATCAGCAACGGAAACCGGTACATCGGTACGGTTGGCTTTCGGATCATGTTCTA
CATCTGTAACGGCTAACGCAAGCCCTATACTTCAGGATGTGAAGGATATTAAATgtgatctctagtttagctttag
tgttccacctgtccataatctgtctttattggtagGACATCGCAAATGCATTGTCGTCTATGACTCAAGTGA
CGTTGATCTAACAGCTGGAAAAGTTAACTACAAACATCGCACGGATTCAAGCTTGTCTGCCAAGCTTAAC
CTGTATCACCGAAAACAGCGGAAGAATTCAAGgtaagcctgtggaaatatgactaaaaagactgctaacaaacga
40 ctctcccaaattgttagCCCCGTTTACAATCGATGAGTGGAGTGTTTTGGCTGATGATCTTCTCATGTTGTGGC
TATCTTCAACCATCATCATGTTGGTTATGTCGCCAGGCTCGACAATGAACGTGCTCTGCCCGTGGCTGGTCC
CGCTGGACTATCACCATTGGACATAGTGTACGGTGTACTGTTCTGTATCAAACCATTTGAATCGTCATGAGCGCA
ACGTACAACCTCTGACGGATAACCAAGTTTCCGGCTTGTGCTACACATAAAATGGACAAATTGTGCCGCTTGTAG
TATGGTTAAAAGgtgagttacggcactacttgccctcagtaaggacaggagttgtttccgttatgatattcatt
45 ttatcagCTTGGACATGACGTCCCTCCGAACGCCAATCGTCGAACGGATGGAAAGCAAGAGATGGAAAGC
GCATCGACCATCACTCCAAAGTGTACGGTACGATGTACGCTAAAGTAACGGAGTGTGCTGTTTACAAGGACATC
TAAAGgtacgaattggccaatttattgtcattttaaaagcttgcaccaactttcacagcttcggcgatgaagt
gcaggacatttccaagGATCTATCTTCCGCAAGTATGCGCGTCTGTAATTATCATTTGTATGACACTGCTGCAAC
TACCGGGGCCATGTTACGATgGCCGATCTGCTGGGCTGTGGGTCTATTGCTAGTAAAGACATCGCAAGTGT
50 TTTCTGTACGTACGGAAATGAAATCTCTATACGgtagttggacacgttagagaattaatgttggaaata
tcaataccaaataatgtatgtttcgttacagACGGATAAAATTACAGAGTTGTIGGGTTTCGAACACTACTCAAC
TTCGATAAGCGTACCGCAAGCAATGATATTCTGCAAATgtgagatagcggtgtatttgctcagtaca

ttaaatacgttctat~~tt~~cag **GACTCTTAAAGATGTTCACATCAAGGTGGAAAGTGTC**TCAAGGTTACGCTAAAT
CTTCACACATTTCAGgtatgtaatttatgctggtagtttagcttgaataagctacaaactttgaaagttaattt
caatctgttttagatTAAGCAGCTATCGTACTCCATCTGGCGTACTTCAGACCATGGAATCAGAGTAATGGCG
5 CTAATATACCTTAATGTTGAAATTATATTTGTTAGATTTATGCATAAAGTAaTaTTAATTATACATCAAACGT
AAGCCGCTaGTTTCAATTAGCCTTCCAAAATTATCAAATTGATTGATTGAGCTTCAAGGTTACGAGTTTCAGGAATT
TAATCTGATAGGATATCTGTTATCCAATAGAGGTGTGGAAGCGTCCCAAGCCATTGTTGATAGTTATAGCA
CCGTCGAGCAGTTGATCGCTGTGATCGCTAGGCGCACCTGATTTATCTTATCTGCACCTGTTATGCCAAGGGCG
10 CTTTCACACGTTCACACAATATAATGCACATGTATAATGCATTCTACTTAGCATTGTTACATATAATACC
AAAATTATGCATTTTATTCTCACGCAACGATTAGAGGATGACTTcACAAAGGTCCATCTAGGGTAGGAGGTATAC
AATTATACCTCTAAAATCTCACAGCAATATGAGAAACAAAGGATACCAAGCATAACCCTTTTACTTGACAATT
TCATTTGATTATGTAATAAAGCACTGCaCGTCGACTTCCTAAAAA

Figure 3a continued

15

.....

Figure 3b

Anopheles gambiae odorant receptor 1 amino acid sequence (SEQ ID NO: 4)

5

MKKDSFFKMLNKHRWILCLWPPEDTDQATRNRYIAYGWALRIMFLHLYALTQALYFKDVKDIND
IANALFVLMQTQVTLIYKLEKFNYNIARIQACLRKLNCTLHYHPKQREFSPVLQMSGVFWLMIFLM
FVAIFTIIMWVMSPAFDNERRLPVPAWFVVDYHHSDIVYGVLFLYQTIGIVMSATYNFSTDPMFSG
LMLHINGQIVRLGSMVKLGHDVPPERQLVATDAEWKEMRKRIDHHSKVYGTMYAKVTECVLF
10 HKDILRIYLRAASMRVCNYHLYDTAATTGGDVTMADLLGCGVYLLVKTTSQVFIFCYVGNEISYTDKF
TEFVGFSNYFKFDKRTSQAMIFFLQMTLKDVKVHKGSLKVTLNHLHTFLQIMKLSYSYLAVLQSM
ESEZ

15

File Edit View Insert Format Tools Window Help

Figure 4a

Anopheles gambiae odorant receptor 2 genomic sequence (SEQ ID NO: 10)

5

Features:

- 1) Presumed Untranslated 5' and 3' regions are underlined.
 - 2) Potential TATA box transcription initiation signal is double underlined.
 - 3) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
 - 4) Introns are tentatively assigned and are shown in lower case.
 - 5) Exons are **highlighted**.

ttccagtaatccataataaaaaataataaaagtaataatagtaaatagtaataattccagtaactgttagtaataac
 acaataatctctaagaattaaaattgcattttgtaatgaaatatgttattttcaggcatttagtttagatatttagatatttaaagcatttgc
 5 ttcattactacaaaaaagcaaatttatgagtgaattacttcagttctctaaacgcctatgttatgcattacat
 aacaatagctcttttattgcattttccttagtaatctaaatccaatctcttcccttgcagATTAAA
 GTCGGCAACGTGTACCCGATGACGTGAAATCTTCAAAATTGCTAACGTGTCTACTCCCTATTTCACACTGCT
 GCGCCGAGTGTACAACTAACCTAACCGGTAAACAAACAAAAATCCCTCATCACTATGCAAAGACAGCAAGCAGGCC
 GATCATCAAACACCATTAGCAGCCACAAAGTTACCAGCCGTTATCCACGGGATTGGTGAAAGTTATTGCACTG
 AAGCTCTTCACCCAAATTTCATGGAGGTTCCCTCTCAACCAACCCATTGAAGCGAATAAAAGTATCAGCAACCAG
 10 GCGACGGTAAAAAACGCTGCATTATTGTGCTTGTCTCAGCATTCCAGCGAATGACTCTTAAACTTTCCATTCAA
 AGTCGCGATGTCACGATA CGGAGCGGTGTTGTTGATCCCGAGTGCACTCGCAAGCCGGTATGGTCCGGT
 GGAAATGCACAGATCGACACAGCGATAGATAATCGTTGTCGCTAAATGGGAGGGAAAAAGTAAGCTGCCAGCT
 ACTTCATTTCCATGTTAATTGAAACTCAAGCCAACGAACATGCGAGAACCCGGTTGGTGTCTCCGCTCCGGGA
 15 AAGGTCCTGCTCCGGGCATGGATTCTTCCCCCTCCGGTGGTTGGGTATTGTTAGGTTTATTTACAAA
 TTCATATCCTTCCGCTTCCGCATCAGCCACCCGGTGGTGCGCCAGACAGATGTGCGCGGGCAACAAACTATGC
 ACGAACATGGCCAACAAACACAGCTTCTATCTCATCTCTGTCGCACTGTCGCTTCCCGCTGCGTTGCTGT
 GTACTATCATTGTTTAGCCACGGTTTACTCTAATTCCATTGCAACCGCAAAAGGCTCATCCTTGCTCGTT
 CCGGTGCAACTTCGACAAGCGATGGTGGGATACGAACAAAAACCAACTACTCCACCCACTACTACTACTG
 CCACCAACCACTAACAAACACTACACTGGTTGGGAGCTTGCAGACCCACAAGCAAACAGATAAGCTAGCTAGCT
 20 GCTGTGTGCGCTCGAGTCAGCGACGGTACAAGGTTAACCGTACAAGCAACTCCGGACCGATCCAAAACACTG
 ACAAGGCACGGGGCGCATCCGGCAGTACGGTCGGAAAACATGGAAATGTTAATTAAACTGTAATTGTCATCGC
 TGCTACAAGTTGTGACACAGGGAGAGAGAGACAGAGCGCGCCGATGGTGTAGGGTGTAAAGATAGATAAGGAA
 AAGAGCGAGAACATTGGTACGATTGGTGTGGTTAGCAAATTGATTTGACTTCACTGATTTGAGTGCAAATTAAATGC
 ATCGAAAATTGGCATTCAAGGTAAGTTGCTCGTGGACGGATCCCCGGGCTGCAGGAATTGATATCAAGCTTAT
 25 CGATACCGTCGACCTCGAGGGGGGGCCGGTACCCAGCTTGTCCCTTAGTGGAA

Figure 4a continued

Figure 4b

Anopheles gambiae odorant receptor 2 amino acid sequence (SEQ ID NO: 6)

5

MLIEECPIIGVNVRVWLFWSYLRPRLSRFLVG CIPVAVLNVFQFLKLYSSWGDMSELIINGYFTV
LYFNLVLRSTSFLVINRKFETFFEGVAAEYALLEKNDDIRPVLERYTRRGRLMSISNLWLGA FISA
CFVTYPLFVPGRGLPYGV TIPGVDVLATPTYQVV FVLQVYLTFPACC MYI PFTSFYATCTLFALVQI
AALKQRLGRLGRHSGT MASTGHSAGTLFAELKECLKYHKQIIQYVHD LNSLVTHLCLLEFLSFGM
10 MLCALLFLLSISNQLAQMIMIGSYIFMILSQMF AFYWHANEVLEASLGIGDAIYNGAWPD FEEPIR
KRLILIIARAQPTDGGKIKVGNVYPMTLEMFKLLNVSYFTLLRRVYN

15

Figure 5a

Anopheles gambiae odorant receptor 3 genomic sequence (SEQ ID NO: 11)

5

Features:

- 1) Presumed Untranslated 5' and 3' regions are underlined.
- 2) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
- 3) Introns are tentatively assigned and are shown in lower case.
- 10 4) Exons are **highlighted**.

AAGCAGAACACATCAAGAAGCAATTAGGTGTCTGACGGTAGCAAGTAGTTCGGAGGGAGAATAAAATAGATGCC
TTCTGAGCGGCCTTCGICTCATTACTTCCCTCGGAACCTCTCAAGACAAACGCAAGATGGTACTGCCAAAATTAAAGG
ATCAAACACCACTGATGCCGTTTCGCTGCAAATTCAAACCATGCCGACTGTTGTTCTACCCAAAGTGTGTTGCCAGGGTAC
CGTTTCTATCTCATCTTCTACTTCTGCCGATGGTGTCTACCCAAAGTGTGTTGCCAGGGTAC
GGTTGCCGTACCGGGACGGCGAGCTGATGTTGAATGAAACGCATTCTGCCATGCTAAATGTTTCTCTTCAAC
GCCACAACTACGAGCGATTGGTGCATCAGCTGCAGGATCTGGCAGCTCTAGgtgagtatgcagccaatcgattgttc
caaacccttcgcacatcctcgtaacactgctacactttcagTCCTCCAAGACCTACCCGAGAGAGCTGGGAGACATAC
CTCATCTCAGTGAACCGACGGGTCGATCGGTTCTCAAAAATTACTGCTGCTGTCACTTTTCATGCCAACGTTCTT
TIGGTTCATGCCCTGCTGGACGGACCTATTCGGCTACTTTGCTGCGCAACAGCACGGAACCGGTGGAGCACGTGT
TGCAGGTGAAGGAAAGACCTGACTCTCTGAACATTGGACTTCGATGGCAGCTATACTGTTTATGTTGCCATTATG
TGGCCCACCGATCTATACGCTGGGTTTACCGGTTGCCACAAAGCTGCTGACCATTTCAGCAATGTTAAGTACTGTT
GCCCATGCTGAAGCTCGTGCACCTCGGAATCCACTGTCTAGGGAGAGTACGGCAAGACCGAGCGAAAGGAGCTGA
ACGAGATTATTCATGCATCAGGGTACTCAAGtaagtaaaattcaaaattgaaagtttcgagggataactttag
tgtgtctgaccctgacatcctagCTGGTGTCTCTGGAGACGACATTGGCTGGTATTTTCTGCACTTC
ATTCACTCTACAAATGATCTGGTGCACTCTCATCTCTACATAGGGTGACGtaatagcattttcgtcatttcgta
gccttattcaatccattttgtaacactgtgaattcccccaggGTTCAGCTCGACGGTAGCGAATGTATGTTCCAG
ATCATTTGCTGACGGTGGAAACTTACGGTACGGCTACTCTGGAAACAGATCTAACCAACCGAGGTGCTTGGtacc
ctttggatgaagcttcaaaaagtaattccaaattctgtttcgattttcccctttccactagAGCTATGGCTTG
CCGGGCCATTACGATAGGGAGCTACAAGTTCCATTTCGATGCCGAAACTTCCGACTGCTACTGCAACGA
TCCCCAAAACCGCTGCCGTAACGGGGAAACTTCCGCTTCCGCAATGTCGGCCAGTTTCCGAAAGgtAACattaat
tacagttgaaaattctgaagaatgcacatcttacttgccttacttggttccagATGCTCAAGATGTCCTTATTCA
TTAGCTACTCTGCAAGGAGCACTTTACGAGCTGTTCCACCCCTGGAAATGGCCTTTCGCACTGCTTCTGT
TTGTTGGACGCACGCAGCACCGAGAGCGCCCTGCACGCACGTACTGATTTGGTACTTTGACGTTGCACCTTG
ACAGCTGAAGGACAGGGTACAATTTCGCTGCTTATTACGCGCAGCGATTGGATACGAAACATTGGCCACAAG
TTCTACGATTTAGCGTTATTACTGTTGCTAGCAGCTTTTCCCaCAATAAACACACACAATAACGTACCGACAG
TATTCTTTCATTGTAGGATAGAGAAGCCGCCAGCAGCCAAACGCGCCGAAACGAAAGGCCACCC
GGGGAAAAACACGGAGCAAACGAGAACAGAACGAGTAACAAACAAACGGCCGAACAACACGGTGCCGGAA
ACGA

40

Figure 5b

Anopheles gambiae odorant receptor 3 amino acid sequence (SEQ ID NO: 8)

5

MPSERLRLITSFGTPQDKRTMVLPLKDEAVMPFLLQIQTIAGLWGDRSQRYRFYLIFSYFCAMV
VLPKVLFGYPDLEVAVRGTAELMFESNAFFGMLMFSQRDNYERLVHQLQDLAALVLQDLPTEL
GEYLISVNRRVDRFSKIYCCCHFSMATFWFMPVWTTYSAYFAVRNSTEPVEHVLHLEEELYFLN
IRTSMAHYTFYVAIMWPTIYTLGFTGGTKLLTIFSNVKYCSAMLKLVALRIHCLARVAQDRAEKEL
10 NEIISMHQRVLNCVFLLETTFRWVFFVQFIQCTMIWCSLILYIAVTGFSSTVANVCVQIILVTVETY
GYGYFGTDLTTEVLWSYGVLAIIYDSEWYKFSISMRRKLRLQQRSQKPLGVTAGKFRFVNVAQF
GKMLKMSYSFYVVLKEQF

15

¶. 0 0 5 6 4 0 5 6 0 0 4 0 5 6 0 0

Figure 6a

Anopheles gambiae odorant receptor 4 genomic sequence (SEQ ID NO: 12)

5

Features:

- 1) Putative Start (ATG) and Stop (TAA) codons are in **BOLD**.
- 2) Introns are tentatively assigned and are shown in lower case.

10 GGGGAACCCCCACCGACCAGACGGAAAGCTAACGATGTCAATTGAATAGTCATTAGT
AGCGTTTGCTCGCAAACGAACTAACCCTTGACTTTAAGTTCACTACCGTGAGGACAAAAA
TCAATAAAATTAAATCGAGACCGTTGATGAGCAAAAGAAAAAAAATTTTACTGATTTCATTT
CGTCCATCGACTACATAATCATAATTATGCCACATTATTATAAGTTTGATCATTTTA
AACACACAAAAATGCATCCTTCGAATATTAGTCAGGTTGTATCAACA**ATGAAGTTGA**ACTGT
15 TTCAAAAATATTCTCCCCGGACACGGCTTATCCTCGTCAAGGCTTGTACATCGTGGGC
ATGAATGGGGCAGGATTCGGTCGCAATTGAGTTGGCATTTCGTTCTATTAATCTT
TCTTGTAAACCGCCACTAACGGCGGGTACACCGATGGTCACCAGCGTACGCACCAGTGTG
GAATTCTGTTAATTGCAATATTACGGCGGCAGTATGTTCTTGCTACGATGTGGCCACTT
CCAAGCGTTCATCCAGGA**ACTGAAGAGC**TTTGGGtaatattaaattaaaattgcgttattgcat
20 catcatttttcatttgcat**TATGCTCACATTG**TACAGACTAAAGTATAAGCTGACCCGGTTAACCGTC
GAGCGGATATTATGCCAAAGTGC**AAACGAC**CTGCATGGGTGCTGTAACGCCTTACTGGAT
TGCACCGATACTTCATCTGCGC**ACTACTACAGGTG**CACCAATTCCACCGAACCCGTGCGG
TTTGTGCAACATTAGAGGTGAAGTTCTATTGGCTCGAGAATCGCACCTCAGTCGAGGACTACAT
AACCTCGTGTGATCATGCTACCCGTCGTGGTTATGTGTGGTTACGTATGCAATTGAAGGTGA
25 TGACCATCTGCTGCAGCATTGGACACTGTACACTGTACACCAGGATGACTATAGAGATGGTAGA
GCAGTTGGAAAGCATGGCATCAGCGGAACGAAC**TG**CCAGCGCCATACGCAACGTGGGGCAGAT
GCACAGTGGTTACTGAAATGCATTAGGCTTGAACACGTCAATCCGATCGATGCTGATGCTGC
AGTGGTTGACCTCGTGTAAACTGGAGCATTCTCTCATCTATCTAACGAACGTGgttagtttgcatt
30 gtttggaaatccaaaaacaaaaagatggctataattgaacttttattacagGGCATCTCGCTACAA**CTGGT**TACCGTGGT
GGTAATGTTTTCTGCCACTGCGGAAACTTCTGTATTGTTACTTGGGACGCGGCTTGC**GA**
CACACAGCAGCTGCTGGAGCACGCACTCTATGCTACACGGTGGTACA**ACTACCC**AATAGCCTT
TCGCAGCAGCATTAGGATGATGTTGAGACAGTCGAAAGGCATGCACACATAACGGTGGGAAG
TTTTTCGCTTAATTGGAAAGAATTAGCAGGATTGTCAACTTATCCTACTCTGCTTACGTCGT
ACTTAAGGATGTAATAAAAGATGGATGTACAGTGAATGTTTTTTGGCTGGCAACGAATGA
35 AGTTTCCGAATCTATATTAGATCTAGAATTAACTAGATGTCAATAATGATCTTGGCCATGA
CCGGTCTGGTTTGGAAACCAATTCTCAAAACAAATTGAACTTAGGGCGAGGCATGAAATGTC
CCAAGAACCTATCCAAGTCTGGAACTACATATTACGAATCTATCCATTATTGCCTCGGA**ACT**
GGTTGGTGCTAAATATTGCTCAAATGTTGGCTCTGGACCTATCCAGACAAAGATCTCAATT
40 TTCCTACCACTGGA**ACTG**ATTAATTGATGTAGGAAGTCATGGAGGTGTCAGGGAGAATTAAA
CACTAATGTC**CAACTC**ATTATTCAAGGGCAATTCTATTTTTATGCCC**CTACGGATTG**ATAC
GTATGTATTACTCCATTCTGGACTTGTCTTATTCTGCTGCTGATTGGACGTGAAATGTTGA
GAAAAAGATTCTTATTGAGTGATA**CAGAGC**TTAAATACTCTACGTTGTTGCTATTAA
GTATGCCAGGCTAATCACA**ATCG**CTACTAATGAACAGAATCTCTAATTAAACCC**TTG**CAT
TGATAGTGTCAATGTC**ATGTC**GAGATAATTGA**ACTG**CAACG**AtAC**CTAC**CTAA**ACGGAGCAG
45 AACACATCAAGAAGCAATTAGGTGTGTCGTACGTTAGCAAGTAGTTCGCGAGGAGGAATAAAAT
AG

50

Figure 6b

Anopheles gambiae odorant receptor 4 amino acid sequence (SEQ ID NO: 14)

5

MKFELFQKYSSPDVTLSFVLRLLHIVGMNGAGFRSRIRVGGIFLFYLFLVIPPLTGGYTDGHQRVR
TSVEFLFNCNIYGGSMFFAYDVATFQAFIQELKSLVCSHSYRLKYKLTRFNRRADIAKVQTTC
MGAVTLFYWIAPIPSICAHYYRSTNSTEPVRFVQHLEVKFYWLENRTSVEDYITFVLIMLPVVVMC
GYVCNLKVMTICCSIGHCTLYTRMTIEMVEQLESMASAERTASAIRNVGQMHSGLLCIRLLNTSI
RSMLMLQWLTCVLNWSISLIYLTNVGISLQSVTVVMMFFLATAETFLYCLLGTRLATQQQLLEHAL
YATRWYNYPIAFRSSIRMMLRQSQRHAHTVGKFFRVNLEFSRIVNLSYSAYVVLKDVIKMDVQ
NVSYSYFTLLRRVYN

15

卷之三

Figure 7

ANOPHELES GAMBIAE

Preferred DNA Codons

Amino Acids			Preferred Codons				
Alanine	Ala	A	GCC	GCG	GCT	GCA	
Cysteine	Cys	C	TGC	TGT			
Aspartic acid	Asp	D	GAC	GAT			
Glutamic acid	Glu	E	GAG	GAA			
Phenylalanine	Phe	F	TTC	TTT			
Glycine	Gly	G	GGC	GGT	GGA	GGG	
Histidine	His	H	CAC	CAT			
Isoleucine	Ile	I	ATC	ATT	ATA		
Lysine	Lys	K	AAG	AAA			
Leucine	Leu	L	CTG	CTC	TTG	CTT	CTA TTA
Methionine	Met	M	ATG				
Asparagine	Asn	N	AAC	AAT			
Proline	Pro	P	CCG	CCC	CCA	CCT	
Glutamine	Gln	Q	CAG	CAA			
Arginine	Arg	R	CGC	CGG	CGT	CGA AGA AGG	
Serine	Ser	S	TCG	AGC	TCC	AGT TCT	TCA
Threonine	Thr	T	ACG	ACC	ACT	ACA	
Valine	Val	V	GTG	GTC	GTT	GTA	
Tryptophan	Trp	W	TGG				
Tyrosine	Tyr	Y	TAC	TAT			

[http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=Anopheles+gambiae+\[gbinv\]](http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=Anopheles+gambiae+[gbinv])

Figure 8

Name	SEQ ID NO	FIG. Reference
Arrestin 1 (cDNA)	SEQ ID NO: 1	Figure 1
Arrestin 1 (polypeptide)	SEQ ID NO: 2	Figure 2
Odorant Receptor 1 (cDNA)	SEQ ID NO: 3	—
Odorant Receptor 1 (polypeptide)	SEQ ID NO: 4	Figure 3b
Odorant Receptor 2 (cDNA)	SEQ ID NO: 5	—
Odorant Receptor 2 (polypeptide)	SEQ ID NO: 6	Figure 4b
Odorant Receptor 3 (cDNA)	SEQ ID NO: 7	—
Odorant Receptor 3 (polypeptide)	SEQ ID NO: 8	Figure 5b
Odorant Receptor 4 (cDNA)	SEQ ID NO: 13	—
Odorant Receptor 4 (polypeptide)	SEQ ID NO: 14	Figure 6b
Odorant Receptor 5 (cDNA)	SEQ ID NO: 15	—
Odorant Receptor 5 (polypeptide)	SEQ ID NO: 16	Figure 9b
Odorant Receptor 6 (cDNA)	SEQ ID NO: 17	—
Odorant Receptor 6 (polypeptide)	SEQ ID NO: 18	Figure 10b
Odorant Receptor 7 (cDNA)	SEQ ID NO: 19	—
Odorant Receptor 7 (polypeptide)	SEQ ID NO: 20	Figure 11b

Figure 9a

Anopheles gambiae odorant receptor 5 genomic sequence (SEQ ID NO: 21)

5

Predicted Exons: *ITALICIZED, UNDERLINED AND HIGHLIGHTED.*
Introns: lowercase.

10 tctagacttgaaccatgacggcatttattgagtgcgttcgactgtaccacgggaccacccgttatcaactatcactatt
aattaattataatatgc~~ttt~~ttagc~~gt~~catc~~g~~c~~cc~~ttttgttctctggatatacttaa~~tt~~tgattatcaagatagaa
caacaacttgc~~tt~~taataatcattac~~gt~~acc~~tt~~aatcaac~~ct~~gtcatcaaggagtttgc~~g~~aa~~g~~caaaaatccgattgtct
gat~~ttt~~gttcttgattccatcgattc~~tt~~act~~gt~~ttctg~~ca~~aa~~at~~gt~~cc~~aa~~ta~~at~~ac~~gg~~ca~~at~~gt~~tc~~tt~~at~~cg~~at~~gt~~ga~~at~~ca~~ac~~at
cacattgtttgcatttgc~~ttt~~ttgc~~gt~~g~~ca~~aa~~at~~at~~gt~~ttatt~~tg~~caa~~a~~agaagg~~ca~~ag~~gt~~taag~~ag~~gtaa~~at~~aca~~at~~tc~~gt~~tg
15 tccatttttgc~~cc~~acc~~ag~~gt~~gt~~gc~~ca~~ga~~ac~~cc~~gt~~gc~~tt~~tt~~ag~~t~~c~~c~~t~~ga~~a~~ta~~c~~at~~c~~cg~~ac~~ca~~g~~gt~~c~~at~~c~~tc~~ATGG~~
~~TGCTACCGAAGCTGTCCGAACCGTACGCCGTGATGCCGCTTCTACTACGCCCTGCAGCG~~
~~TTTCGTTGGGCTGTGGGGTGAACGACGCTATCGCTACAAGTTCCGGTTGGCATTTTA~~
~~AGCTTCTGCTGCTAGTAGTTATTCCGAAGGTTGCCCTCGGCTATCCAGATTAGAGAC~~
~~AATGGTTCCGCGAACAGCTGAGCTGATTTGAATGGAACGTACTGTTGGATGTTG~~
20 ~~CTGTTTCTCTCAAGCTAGACGACTATGATGATCTGGTGTACCGGTACAAGGACATATC~~
~~AAAGATTGgtgcgtataatgattgataaaagaac~~tt~~gag~~ca~~act~~ct~~at~~cc~~tt~~ca~~ag~~C~~TTTCCGTAAGGAC~~
~~GTTCCCTCGCAGATGGGC~~ACT~~ATCTGGTACGCATCAATCATCGTATCGATCGGTTTC~~
~~CAAGATCTACTGCTGCAGCCATCTGTGTTGCCATCTTCTACTGGGTGGCTCCTTCG~~
~~CCAGCACCTACCTAGCGTACCTGGGGGCACGAAACAGATCCGTCCCCGGTCGAACATGT~~
25 ~~GCTACACCTGGAGGGAGGACTGTGACTGGTTACACACCCGGTCTCGCTGGTAGATTAC~~
~~TCCATATTCA~~CCGOC~~CATCATGCTGCC~~T~~ACAATCTTTATGCTAGCGTACTTCGGTGGACT~~
~~AAAGCTGCTAACCATCTTCAGCAACGTGAAGTACTGTTGGCAATGCTCAGGCTTGTG~~
~~GCGATGAGAATCCAGTTCATGGACC~~GG~~CTGGACGAGCGCGAAGCGAAAAGGA~~ACT~~GA~~
~~TCGAAAATCATCGTCATGCATCAGAA~~GGCG~~CTAA~~Ag~~taagg~~t~~tgc~~cc~~gtat~~tt~~tg~~gg~~at~~aa~~at~~ac~~attt~~
30 ct~~ag~~ctgc~~tt~~c~~ag~~ ATGTGTGGAGCTGTTGGAAATCATCTT~~CGG~~TGGGTTTTCTGGGACAG
~~TTCATACAGTGC~~G~~TAATGATCTGGTGCAGCTGGTTCTGTACGTGCGCCGTTACG~~gt~~ta~~act~~aa~~agg~~ac~~ct~~gt~~at~~gt~~ct~~gc~~c~~ac~~c~~ac~~tt~~ca~~ct~~gt~~gt~~tt~~gt~~tt~~gt~~ca~~ct~~tt~~cc~~ca~~g~~G~~GTCTCAGC~~AC~~AAAAG
~~CGGCAAACGTGGGTACTGTTATACTGCTAACAGTGGAAACCTACGGATTCTGCTA~~
~~CTTGCCAGT~~G~~ATCTTACCTCGGAGGCCAAGTTGTTATTGCTGAGttc~~ag~~ttactttcc~~gt~~ttcccc~~
35 tctaaccgtaccacttgtaccatttgtt~~g~~ag~~ac~~ag~~ag~~gt~~gt~~g~~cg~~ct~~ag~~ CACGTGCTGC~~GT~~ACCGTAGCCTCTGG
~~TATCGCCGTTCGGTT~~CG~~ATTCAACGGAA~~AG~~CTTCGAATGGTACTGCAGCGT~~GCCC~~CAGA~~
~~AACCGGT~~CG~~GCATCTCGGCTGGGAAGTTGCTTCG~~TC~~CGACATTGAGCAGTTGGCAA~~
~~Tgtatggggagacc~~tt~~caactgtggcaagaa~~ag~~tttcttattatgc~~at~~tttaatttacag~~AGGG~~AAAACATCA~~
~~TACTCGTTCTACATCGTTCTGAAGGATCAATTAA~~agg~~gaact~~ccc~~c~~ac~~c~~g~~gacc~~ac~~gac~~gg~~aa~~
40 agctaac~~cg~~at~~gt~~gc~~aa~~t~~gt~~ga~~at~~gt~~ct~~at~~gt~~tg~~ttt~~gc~~ca~~aa~~ac~~gc~~aa~~act~~ac~~cc~~tt~~tg~~at~~tt~~ta~~ag~~ttt~~tc~~at~~tt~~cg~~gt~~gg~~ag~~g~~
gacaaaaat~~ca~~ataaa~~at~~tc~~g~~ag~~ac~~cc~~gt~~tg~~at~~g~~ag~~ca~~aa~~ag~~aaaaaa~~at~~ttt~~act~~g~~at~~ttt~~c~~at~~tc~~at~~cg~~act~~a
cataatcataattatgc~~cc~~ac~~at~~tttattataag~~ttt~~tg

45

Figure 9b

Anopheles gambiae odorant receptor 5 amino acid sequence (SEQ ID NO: 16)

5

MVLPKLSEPYAVMPLLLRLQRFVGLWGERRYRYKFRLAFLSFCLLVVIPKVAFGYPDLE
TMVRGTAELIFEWNVLFGMILLFSLKLDDYDDLVYRYKDISKIAFRKDVPSQMGDYLVRI
NHRIDRFSKIYCCSHLCLAIFYWVAPSSSTYLAYLGARNRSVPVEHVLHLEEELYWFHTR

10

VSLVDYSIFTAIMLPTIFMLAYFGGLKLLTIFSNVKYCSAMLRLVAMRIQFMDRLDEREA
EKELIEIIVMHQKALKCVELLEIIFRVVFLGQFIQCVMIWCSLVLYVAVTGLSTKAANVG
VLFILLTVETYGF CYFGSDLTSEASCYSLTRAAYGSLWYRRSVSIQRKLRMVLQRAQKP
VGISAGKFCFVDIEQFGNMAKTSYSFYIVLKDF

15

File Edit View Insert Cell Tools Help

Figure 10a

Anopheles gambiae odorant receptor 6 partial genomic sequence (SEQ ID NO: 22)

- 5 These are the predicted last three exons of another candidate *Anopheles gambiae* odorant receptor.

Predicted Exons: *ITALICIZED*, UNDERLINED AND **HIGHLIGHTED**.

Introns: lowercase.

10

aacaccatttatcgcaaaatttagtattaccgttggaaagcggttcctcggcttttcactctctctgtctctta
ttgatccgtatgcgcgcgtctataggtag*TTATGCTTACCGGATGTTGGATCCGCACGTGCTTT*
TCCGCATACGCCAGTGCACACTTGATGGCGGTGGTATGACGTCTGCTGCCACCGTT
15 *TTCTGCTCGTGAGTCAGACCTTTCAATTCTGCAATATCCTGTTCTTCCCCAACCCC*
ACAGACGGTTAGACGGATATATGCTGGTAAAGTTGTCCTCTCATGCTGTGCTTCTG
ATCGAGCTGCTGATGCTGTGCGTAGGGATATTGTTGGAATCGgttaaggcaccaggc
20 *ggtagatgagcgagtcgcgagaattgaagctttgtttaaaacacatcagaag**CCTTGGGGTGATTGATGCCGCT*
TACGGTTCCGAATGGTACCGGAAAGGGTGGTGGCGTCCATCGATCCGTGCTGCAA
25 *TTATACACCCGAGCCAGCAGTCCGTCAACTGACCCGATGGAAAATTGGCCCATCCAA*
ATGAGTACTTCAGTCAGtgagttgccaattgatgcgtttgcgttaatattcagtaagagtgcgtctttcccttag
*ATCCTGCAAGCTTCTGGTCTACTTTACCCCTCTGAAGACCGTCTACGGGAATA*Agtaa
gcgcgagagagagagagcagtategttcacccttggatgaatcaatagattctaatacgttgcaccacgaccaagcggttgtcatcaggaccaaacacgtttca
acatttcgttagttgcacaatattgtaccattctatacagcttgcaccacgaccaagcggttgtcatcaggaccaaacacgtttca
caagccgcgtcacctgtggc

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

90

95

100

105

110

115

120

125

130

135

140

145

150

155

160

165

Figure 10b

Anopheles gambiae odorant receptor 6 partial amino acid sequence
(SEQ ID NO: 18)

5

LCLPDVAIAHVLFRIRQCTLGGGDDVCCAPFSARESDLFISCNILFLSRPHRRLDGYML
VKFVLFMLCFLIELLMLCAYGEDIVESPWGDZCRLRLRMVPGRVGGVPSIRAANYTPQP
AVRHTDRMENLAHPNEYFQSDPASFLVLLYPPEDRLRE

10

Figure 11a

Anopheles gambiae odorant receptor 7 genomic sequence (SEQ ID NO: 23)

5

Features

1. Predicted Exons (7): **ALL CAPS, ITALICIZED, UNDERLINED, HIGHLIGHTED**

2. Introns (6): lowercase

3. 5' and 3' sequences: lowercase, dotted underlined

ccggccggcagggtactacgcggtctgacttgcgtgcgtgcctttgtacggcaaacggctacacaagcgaaatcgaaattttcc
tatcacgctgcgttaccagcgcgtgttaggcaaagaatgtcaaaagttcattggcttgcgtgtgtgtgtgtgtgcagtggaa
gcacgggtcatcgtaagggttcgggtgagccagaagttcagatcgaaatctttgtgtgtgtgtgtgcagtggaa
gcattgtgttagtggaaatgtgaaaagaaaagtgtgaaatgtcaagtcacgtccagccgaccaagtacgtcggccctgtgcccaccc
gatgccgaacattgggtgatcaggccagggcaacttgcgttgcgtacgtcaccggccgatactgatccgcaagggtgtac
tcctgggtggacgctgccc **ATGGTGCTGATCCAGTTCTCGCCATCCTCGGCAACCTGGCGACGA**
ACGGCGGACGACGTGAACGAGCTGACCGCCAACACGATCACGACCCIGTTCTCACCGCA
CTCGGTCAACCAAGTTCATCTACTTTCGGTCAACTCGGAGAACTTCTACCGGACGGCTC
GCCATCTGGAACCCAGACCAACACGCACCCGCTGTTGCCGAATCGGACGCCGGTACCG
ATTGCGATTGGCGCTCGCCAAGATCGGAAAGCTGCTGGTGCTGGTGTGGCGATGCCACCCACCGT
CCTGTCGGTTGTCGgtatgtgttatgtgtggccgttggaaagtgtcttgcggcagaacccaaatctactttacgc
ttgactgggtttttttctcggtggagggacgggataaaaatctgaaagaataattgagtcaacccacagggggatgcaag
acatcgccaggcagaggttgggttgcattatcacgcacaccgaatatctcaggttcataaagtttcacccgcgtgaaaaggaa
actccccattccctgttttttttcctctcgataaaattactcatgcgtttcgttttttttttgtgttgcttttttcatc
cctactag **CCTGGGTTACGATAACATTTCGGCGAGAGCGTCAAGACTGTGCTCGATAAG**
GCAACCAACGAGACGTACACGGTGGATAACCCCCGGCTGCCCATCAAGTCTGGTATC
CGTGGAAATGCAATGAGCGGACCGGCGTACATTCTCTTCACTACCAGGTACGGT
GCAGGAATgtcctgcgtcacagtggcagtcgtcgtgagggcaacacggcaaaaaaaaatggactaaaacggttttcacaga
gccaacacattctacagcaattgcataccctcggtggactggcaatgcagtcacacatctcgctaaagttatgcata
tcgagcagaaaaatgtggcggttagggttttgtgataatagtcgtttttgtctctcgcttataaaactctatcaacggaggaaa
tccatttcgctacaatgcctacagctcaagttcaaggtcaatcgagcgggtgggatcaacttttattcatttgcataacgcccc
tcaacaaattctatgttctcaatggcaaagattactgcggcacaatgcggcaacggcaaaaagaaaagcgacgattatga
agatgtccaaaccattggccggcgcgttgcgttgcgtttactgtgtttactgtctgtactttcaggcacaaggaa
atgaaaccagcgcaggcgtttcggtttgcggctgcggactggcaatgcagtcacacatctcgctaaatgcgatattttacgatt
tggatccagtttatgtatgtggcctgcattacagttgcattataccctgtatgttcattgcatttgcattttgtaaagttgtgtgt
cccgtaacgattaattcttcaaaagagattcttcaaaagagattcaaaatgtgtataacaaatgtcaacgaatggaccgtacttgg
agggttgcgaaagtaacgtttaaaatattcatcacaatctctgcggcaacttgcgttgcattttgtgcacaataagttaaaact
gtggcggcagatgtgtcgctgtcgcttcctccctccagcaagctgtggcaaaataatttattccatcatttaatacagccgttgc
cattttaaatggcaatggcaataaaaagcagctaaccatccccattaaaacaaaatgtgttgcgttgcggccatgttatggcgggt
aagtaatggtttaccagtggaaagtgccttccatgtgggtacttcgcgtattttgttatacaagtgcatacagaaaaaaa
ggacaaaatcccttgctatgttcaaggccagctcggtaccgttccggttgcgttgcattaaatgtgttgcataaagttgt
acttcgcgttcaaccaccaatggactttcatgttgcgttgcattaaatgttgcgttgcattttgtgcacaataagttaaaact
tcggcggcgcctcatccccagtttgcgcaccaatattgccttcattaatctgttgcgttgcgttgcggccatgttatggcgggt
cggttgcattatgcacccgcgttgcacgggataatcggtggacggcgttgcgttgcgttgcggccatgttatggcgggttgcacccg

gtaccgcacccatccgtaccgataccggAACAAACGGTGTGCGCGAAAGAATCCGCTAGCAGCCCCACTGGCACGGTATTGCTT
TGGTTCTGTGTTTCTTCCACTGGTTGGGTGCCTGGCGAAGGCTAGCTCGGTACTTCCGGGCCGAATTTCTGCAGCCCAAG
CGGGCGTGCCTGGGGCCAAAAGAAT

5

Figure 11a continued

File Edit View Insert Format Tools Help

Figure 11b

Anopheles gambiae odorant receptor 7 amino acid sequence (SEQ ID NO: 20)

5 MVLIQFFAILGNLATNADDVNELTANTITTLFFFTHSVTKFIYFAVNSENFYRTLAIWNQT
NTHPLFAESDARYHSIALAKMRKLLVLVMATTVLSVVAWVTITFFGESVKTVLDKATN
ETYTVDPRLPIKSWYPWNAMSGPAYIFSFIYQVRWRNGIMRSLMELSASLDTYRPNSSQ
LFRAISAGSKSELIINEEKDPDVKDFDLSGIYSSKADWGAQFRAPSTLQTFDENGRNGNP

10 NGLTRKQEMMVRSAIKYWVERHKHVVRLVSAIGDTYGPALLHMLTSTIKLTLLAYQA
TKIDGVNVYGLTVIGYLCYALAQVFLCIFGNRLIEESSVMKAAYSCHWYDGSEEAKT
FVQIVCQQCQKAMTISGAKFFTSLDLFASVLGAVVTYFMVLVQLK

15